

THE GISMO 2-MILLIMETER DEEP FIELD IN GOODS-N

JOHANNES G. STAGUHN^{1,2}, ATTILA KOVÁCS^{3,4}, RICHARD G. ARENDT^{2,5}, DOMINIC J. BENFORD², ROBERTO DECARLI⁶, ELI DWEK², DALE J. FIXSEN^{2,7}, GENE C. HILTON⁸, KENT D. IRWIN^{8,9}, CHRISTINE A. JHABVALA², ALEXANDER KARIM^{10,11}, SAMUEL LECLERCQ¹², STEPHEN F. MAHER^{2,13}, TIMOTHY M. MILLER², S.HARVEY MOSELEY², ELMER H. SHARP^{2,14}, FABIAN WALTER⁶, EDWARD J. WOLLACK²

Rev. 06-07-14

ABSTRACT

We present deep continuum observations using the GISMO camera at a wavelength of 2 mm centered on the Hubble Deep Field (HDF) in the GOODS-N field. These are the first deep field observations ever obtained at this wavelength. The 1σ sensitivity in the innermost $\sim 4'$ of the $7'$ diameter map is $\sim 135 \mu\text{Jy}/\text{beam}$, a factor of three higher in flux/beam sensitivity than the deepest available SCUBA 850 μm observations, and almost a factor of four higher in flux / beam sensitivity than the combined MAMBO/AzTEC 1.2 mm observations of this region. Our source extraction algorithm identifies 12 sources directly, and another 3 through correlation with known sources at 1.2 mm and 850 μm . Five of the directly detected GISMO sources have counterparts in the MAMBO/AzTEC catalog, and four of those also have SCUBA counterparts. HDF850.1, one of the first blank-field detected submillimeter galaxies, is now detected at 2 mm. The median redshift of all sources with counterparts of known redshifts is $\bar{z} = 2.91 \pm 0.94$. Statistically, the detections are most likely real for 5 of the seven 2 mm sources without shorter wavelength counterparts, while the probability for none of them being real is negligible.

Subject headings: galaxies: evolution – galaxies: high-redshift – galaxies: luminosity function – galaxies: photometry – galaxies: starburst – infrared: galaxies

1. INTRODUCTION

Submillimeter and millimeter observations have revealed the existence of a population of previously unknown high-redshift dust-enshrouded starburst galaxies. Virtually all their stellar UV and optical radiation is absorbed and reradiated by the dust at infrared (IR) wavelengths. They are among the most luminous galaxies in the universe, and their relative contribution to the galaxy number counts and co-moving luminosity density increases with redshift (e.g. Sargent et al. (2012)).

The most massive galaxies are predicted to be at the center of galaxy clusters that reside in the most massive dark matter halos. Surveys that map their distribution with redshift will therefore reveal the epochs of cluster formation in the early universe. For example, follow up observations of two submillimeter galaxies at optical and near-IR wavelengths have shown that they are members of protoclusters that formed at $z \approx 5$ (AzTEC-3: Capak et al. (2011), HDF850.1: Walter et al. (2012)). A survey of dusty starbursts is also essential for determining the obscured cosmic star formation rate at high redshift, and for understanding the formation and evolution of dust in these objects.

The advantage of using (sub)millimeter wavelength observations to search for these objects stems from the fact that starburst galaxies have typical dust temperatures of 35 K, so that their IR spectrum peaks at $\sim 90 \mu\text{m}$. Submm-mm observations therefore trace the Rayleigh-Jeans part of their spectrum, and benefit from the fact that the decrease in flux from high redshift objects is largely offset by the negative K-correction.

Figure 1 depicts the flux of a typical dusty submillimeter galaxy versus redshift (solid lines). This starburst galaxy is characterized by an IR luminosity of $10^{12} L_{\odot}$ and a dust mass of $10^8 M_{\odot}$. The dust was assumed to have a $\kappa(\lambda) \propto \lambda^{-\beta}$ mass absorption coefficient with a spectral index of $\beta = 1.5$, and a temperature of 35 K. An interesting effect at high redshifts is the fact that dust heating by the cosmic microwave background (CMB) becomes comparable to the heating by ambient starlight. When accounting for both sources of heating, the actual dust temperature can be expressed as $T_d = (T_0^{4+\beta} + T_{CMB}^{4+\beta})^{1/(4+\beta)}$, where $T_{CMB} = 2.73(1+z)$ is the CMB temperature at redshift z , and T_0 is the dust

¹ The Henry A. Rowland Department of Physics and Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

² Observational Cosmology Lab, Code 665, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

³ California Institute of Technology 301-17, 1200 E California Blvd, Pasadena, CA 91125

⁴ Institute for Astrophysics, University of Minnesota, 116 Church St SE, Minneapolis, MN 55455, USA

⁵ CRESST, University of Maryland – Baltimore County, Baltimore, MD 21250, USA

⁶ Max-Planck-Institute für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany

⁷ CRESST, University of Maryland, College Park, College Park, MD 20742, USA

⁸ NIST Quantum Devices Group, 325 Broadway Mailcode 817.03, Boulder, CO 80305, USA

⁹ Dept. of Physics, Stanford University, Stanford, CA 94305, USA

¹⁰ Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

¹¹ Argelander-Institut für Astronomie, University of Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

¹² Institut de Radio Astronomie Millimétrique, 300 Rue de la Piscine, 38406 Saint Martin d'Heres, France

¹³ Science Systems and Applications, Inc., 10210 Greenbelt Rd, Suite 600, Lanham, MD 20706, USA

¹⁴ Global Science & Technology, Inc., 7855 Walker Drive, Suite 200, Greenbelt, MD 20770, USA

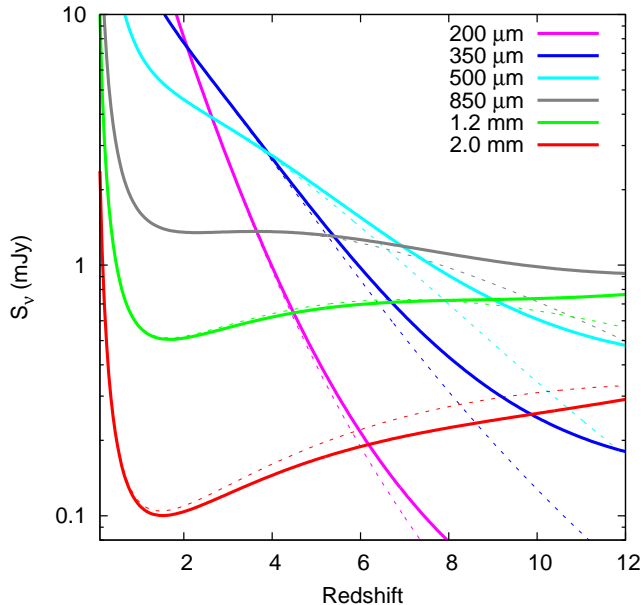


FIG. 1.— Redshift-dependent flux density, measured against the CMB, of a galaxy with fixed dust mass $M_d = 10^8 M_\odot$, dust emissivity index of $\beta=1.5$, i.e. with a FIR luminosity of $L \sim 10^{12} L_\odot$, shown for a variety of wavelengths, based on the multi-temperature empirical dust models of Kovács et al. (2010). To radiate the same luminosity against an increasingly warmer CMB in earlier epochs, the cold dust temperature ($T_0=35$ K) must rise as $T_d^{4+\beta_{\text{eff}}} = T_{\text{CMB}}^{4+\beta_{\text{eff}}} + T_0^{4+\beta_{\text{eff}}}$ (the effective dust emissivity, β_{eff} , is defined in Kovács et al. 2010). For comparison, the dashed lines show the same if the CMB heating is ignored. Note, that the observed flux density at 2 mm wavelengths increases monotonically and steeply as a function of redshift for $z > 1$.

temperature when heated by starlight alone. GISMO (the the Goddard IRAM 2 Millimeter Observer) observations are a differential measurement of the flux from a galaxy against the CMB. The observed galaxy spectrum in such measurement is thus given by:

$$F_\nu(\lambda) = 4\pi M_d \kappa(\lambda) [B_\nu(\lambda, T_d) - B_\nu(\lambda, T_{\text{CMB}})] \quad (1)$$

which cannot be characterized by a single blackbody with a simple $\lambda^{-\beta}$ emissivity law. The dotted lines in the figure show the fluxes that would be measured if the CMB radiation were not present.

The figure shows that the 2 mm fluxes tend to be lower than those at the shorter wavelengths. However, the rising 2 mm flux with redshift provides the strongest bias towards the detection of high redshift galaxies. Furthermore, the atmospheric transmission is higher, and the atmospheric background noise is lower at 2 mm than at shorter wavelengths.

We have developed the GISMO instrument that utilizes a near background-limited detector to fully exploit the advantages of the 2 millimeter window. Here we report the first deep survey conducted with GISMO centered on the Hubble Deep Field North (HDF-N). The Hubble Deep Field North (HDF-N) is one of the best studied regions in the sky. Its sky coverage is one HST WFPC2 pointing, i.e. $\sim 2.5' \times 2.5'$, which is less than the $2' \times 4'$ instantaneous field of view of the GISMO array. The HDF is located in the greater GOODS-N region, which has also been studied in exquisite detail over the last decade

at many wavelengths, X-rays: Brandt (2008), UV: Teplitz et al. (2006), Optical: Giavalisco et al. (2004), optical spectroscopy: Cowie et al. (2004), Near-Infrared: Yan et al. (2006), Mid-Infrared: Rodighiero et al. (2006), Far-Infrared: Borys et al. (2003), Frayer et al. (2006), Perera et al. (2008), Radio: Morrison et al. (2010). In the (sub-)millimeter regime, the HDF and GOODS-N have been studied by the (sub-)millimeter cameras SCUBA, MAMBO, and AzTEC in the past (Hughes et al. (1998), Pope et al. (2005); Greve et al. (2008), Penner et al. (2011)). The currently available data of the full HDF at 1 mm reach $1 - \sigma$ sensitivities of 0.5 mJy/beam (MAMBO observations combined with AzTEC observations, (Penner et al. 2011), the SCUBA “super”-map (Pope et al. 2005) reaches a peak depth of $0.4 \text{ mJy beam}^{-1}$ at $850 \mu\text{m}$, however the sensitivity varies significantly over the observed area in the field.

The paper is organized as follows: We first describe the instrument and its characteristics in §2. In §3 we describe observations and the data reduction. In §4 we describe the source extraction and its results, and present simulations used to characterize the data and to evaluate the completeness and reliability of the extracted sources. §5 presents the 2 mm number counts and the analysis of the properties of select individual sources.

2. THE GISMO 2 MM CAMERA

Continuum observations in the 2 mm atmospheric window have not been astronomically explored from the ground to the same degree as has been done at shorter wavelengths (1 mm or less), except for Sunyaev-Zel’dovich (S-Z) observations with dedicated 6 to 10 m class telescopes (Swetz et al. (2011), Carlstrom et al. (2011), Dobbs et al. (2006)). The reason for this is predominantly of technical nature, in particular the very demanding requirements on the noise performance of a background-limited camera operating in this low opacity atmospheric window. In order to provide background-limited observations in the 2 mm window at a good mountain site such as the the IRAM 30 m telescope on Pico Veleta, the required sensitivity, expressed in Noise Equivalent Power (NEP), for the detectors is $\sim 5 \times 10^{-17} \text{ W } \sqrt{\text{s}}$ (Staguhn et al. 2006), a requirement that is met by our “high” temperature ($T_c = 450 \text{ mK}$) Transition Edge Sensor (TES) detectors. Consequently we have proposed and built a 2 millimeter wavelength bolometer camera, the Goddard-IRAM Superconducting 2 Millimeter Observer (GISMO, Staguhn et al. (2008)) for astronomical observations at the IRAM 30 m telescope on Pico Veleta, Spain (Baars et al. 1987). GISMO uses a compact optical design (Sharp et al. 2008) and uses an 8×16 array of close-packed, high sensitivity TES bolometers with a pixel size of $2 \times 2 \text{ mm}^2$ (Benford et al. 2008), which was built in the Detector Development Laboratory at NASA/GSFC (Allen et al. 2008). The array architecture is based on the Backshort Under Grid (BUG) design (Allen et al. 2006). GISMO’s bandpass is centered on 150 GHz and has a fractional bandwidth of 20%. The superconducting bolometers are read out by SQUID time domain multiplexers from NIST/Boulder (Irwin et al. 2002). This design is scalable to kilopixel size arrays for future ground-based, suborbital and space-based X-ray and far-infrared through millimeter cameras (e.g. Staguhn et al. (2012)).

3. OBSERVATIONS & REDUCTION

The GISMO Deep Field (GDF) observations of the HDF-N were obtained between April 13 and 18, 2011, and on April 11, 12 and 23, 2012. The total integration time was $t \sim 39$ hours, however 2/3 of those observations were obtained with GISMO's lower sensitivity during the Spring 2011 run (see section 3.4). The FWHM of GISMO's beam is $17.5''$.

3.1. Data Reduction

The data were reduced, using CRUSH¹⁵ (Kovács 2008), which is the standard reduction software for the GISMO camera. CRUSH is open-source and available in the public domain. The data reduction tool of CRUSH consists of a highly configurable pipeline, which uses a series of statistical estimators in an iterated scheme to separate the astronomical signals from the bright and variable atmospheric background and various correlated instrumental noise signals. It determines proper noise weights for each sample in the time series, removes glitches, identifies bad pixels and other unusable data, and determines the relevant relative gains. It also applies appropriate filters for $1/f$ -type noise, and other non-white detector noise profiles. For the detection of point sources, the resulting "deep"-mode maps are spatially filtered above $50''$ FWHM to remove spatially-variant atmospheric residuals. The fluxes in each 10-minute scan are corrected for the line-of-sight atmospheric opacities, based on the IRAM radiometer measurements. Point-source fluxes are also corrected scan-wise for the flux-filtering effect of each and every pipeline step, and for the large-scale structure filtering of the final map. As a result, comparison to point-like calibrator sources (e.g. planets and quasars) is straightforward, even if different reduction options are used for these and the science targets.

3.2. Calibration

Mars, Uranus, and Neptune were observed for primary flux calibration. Of those, measurements of Mars cover the widest range of weather conditions. Using the atmospheric transmission model of the Caltech Submillimeter Observatory¹⁶ and the IRAM 30 m Telescope 225 GHz radiometer readings, we obtain excellent calibration in effectively all weather conditions: a 7% rms blind calibration up to $\tau_{225\text{GHz}} \sim 1$ is obtained. Note that any model uncertainties due to the different elevation of Mauna Kea, the site of the CSO, and Pico Veleta, will be very small and therefore irrelevant for the accuracy of derived calibration factors. Figure 2 shows the histogram of the 2 mm line-of-sight opacities for all data.

3.3. Pointing & Astrometric Accuracy

During the GISMO observing runs in 2011 and 2012, we obtained a large number of pointing measurements over the entire sky, from which we derived appropriate pointing models according to Greve et al. (1996). Our pointing models yield $<3''$ rms accuracy in both Az and El directions on all pointing measurements obtained during the two observing runs (424 and 392 individual pointing observations for the 2011 and 2012 observing runs,

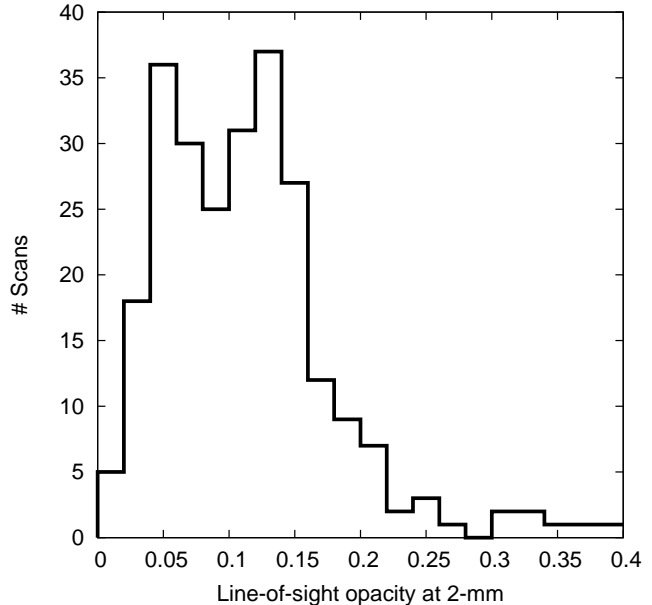


FIG. 2.— Histogram of the 2-mm line of sight opacities for the GDF observations

respectively). Additionally, we frequently checked pointing on nearby quasars during GDF observations. Triggered by a reduction flag, CRUSH will automatically incorporate the measured differential offsets with respect to the pointing model, to further improve pointing accuracy, and to remove most systematic pointing errors in the pointing model, or due to structural deformations of the telescope. The resulting residual pointing errors are expected to be independent and random between independent pointing sessions. Thus, a representative lower bound to the final astrometric accuracy is given by the instantaneous pointing rms ($<4.2''$) divided by the square root of the number of independent pointing sessions spanning the observations. In our case, approximately 30 independent pointing sessions bracket the GDF observations. Therefore, the astrometric accuracy of our map (notwithstanding the inherent positional uncertainties of any detections) could be as low as $0.8''$ rms, or somewhat higher in the presence of systematics errors, which are not eliminated by the use of nearby pointing measurements.

3.4. Instrument Performance

The noise equivalent flux density (NEFD) of measurements during the 2011 run was typically $15\text{--}17 \text{ mJy}\sqrt{\text{s}}$, under most weather conditions. The obtained sensitivity at that time was mainly limited by a neutral density filter with 40% transmission. This filter was needed, since there was a significant amount of THz light scattered into the GISMO beam by the low pass filters, which were positioned very close to the entrance window of the dewar. In early 2012 we mounted a 77 K baffle dewar in front of the GISMO optical entrance window, which reduces the stray light significantly and eliminates the need for a neutral density filter in the instrument (Sharp et al. 2012). As a result of this, the NEFD obtained during the 2012 observing run was typically $10 \text{ mJy}\sqrt{\text{s}}$.

3.5. Noise properties of the beam-smoothed map

¹⁵ <http://www.submm.caltech.edu/~sharc/crush>

¹⁶ <http://www.submm.caltech.edu/cso/weather/atplot.shtml>

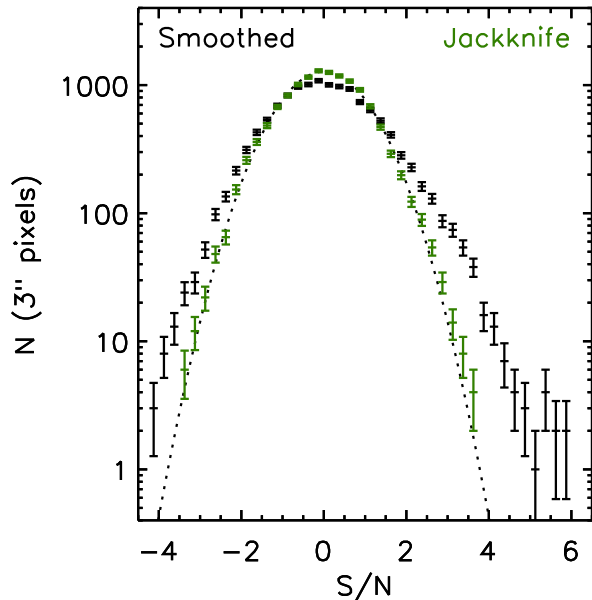


FIG. 3.— Signal to noise histogram for the smoothed and filtered GDF data (black), which was used for the source extraction, and shown in Fig. 5, and for a corresponding time-based random jackknife realization (green). The dotted line shows the expected Gaussian noise distribution, based on the radiometric down-integration of the detector timestream noise, with $\sigma = 1.00$. The close agreement of the jackknifed noise distribution and the Gaussian expectation indicates that our measurement noise is both closely Gaussian in nature, and radiometric. At the same time, the histogram of the regular (non-jackknifed) map exhibits distinct deviations from the Gaussian noise. A symmetric widening is caused by additional noise from unresolved sources (i.e. confusion noise) on top of the radiometric measurement noise, while resolved emission sources cause the asymmetric excess on the positive half of the distribution.

To estimate the noise we randomly multiply each of the individual 10 minute scans by $+1$ or -1 , a method known as “jackknifing”. This eliminates any stationary noise (including sources, and foreground) but retains random noise, including that from the atmosphere. The histogram of the S/N for the jackknifed, beam-smoothed, and filtered GISMO map is shown in Figure 3. The distribution is well fit by a Gaussian with $\sigma = 1.00$. The S/N histogram for the regular (not jackknifed) smoothed and filtered map (also Fig. 3) shows distinct excess deviations from the Gaussian distribution on both extremes of the distribution. When we subtract the 12 detected sources (see section 4.3) from the image, the S/N histogram does approach the expected noise distribution as shown in Figure 4. The subtraction of sources (both real and false) causes this histogram to be truncated at the 2.99σ detection limit. Below this limit both the positive and negative sides of the histogram are more closely Gaussian because the effective smoothed and filtered PSF used for the subtraction has both positive and negative features (main beam and surrounding filter bowl). There is, however, a slight ($\sim 5\%$) symmetric excess remaining in the post-extraction residuals, relative to the jackknifed map. This excess would be consistent with the presence of confusion noise in our map. We cannot, however, quantify the corresponding level of 2-mm confusion noise with any

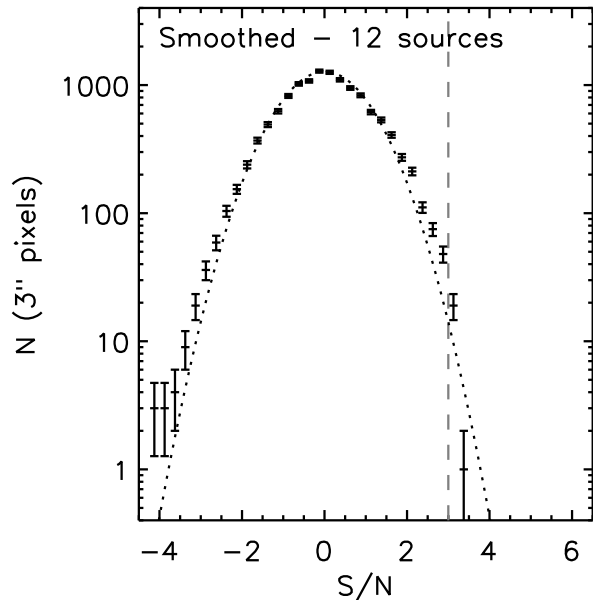


FIG. 4.— Signal to noise histogram of the smoothed and filtered GDF data (Fig. 5), after the 12 blindly identified sources are removed and flagged. The vertical dashed line marks the source extraction threshold, resulting in an abrupt truncation of the histogram above the detection threshold. The dotted line shows the expected Gaussian noise distribution from the jackknifed map with $\sigma = 1.00$. The post-extraction residuals are more closely Gaussian, as expected, albeit still hinting at the presence of further (fainter) resolved and unresolved sources in the field, which manifest as an asymmetric and a symmetric excess, respectively.

precision because the the noise estimation errors are relatively large given the low number of beams in the field.

4. SOURCE EXTRACTION & SIMULATIONS

4.1. Extraction – Method and reliability

The Gaussian nature of noise in our map (section 3.5) is a notable feature of the GISMO data and allows us to provide the strongest possible statistical characterization of our source candidates relying exclusively on formal Gaussian statistics. Because we do not detect any deviation from Gaussian noise, nor see any signs of non-radiometric down-integration in the jackknifed maps, we need not worry about potential troublesome statistical biases that could otherwise result from non-Gaussian, or correlated, noise features.

We extract sources from the beam-smoothed filtered map, which were produced by CRUSH. Beam-smoothing is mathematically equivalent to maximum-likelihood PSF amplitude fitting at every map position (Kovács et al. 2006b). I.e., the PSF-smoothed map value and its noise directly provide the amplitude and uncertainty of a fitted PSF at each map position. The effective filtered PSF for GISMO maps reduced in ‘deep’ mode in CRUSH is accurately modeled as a combination of a $17.5''$ FWHM Gaussian main beam combined with a negative $50''$ FWHM Gaussian bowl such that the combined PSF yield zero integral (signifying that we have no DC sensitivity due to the sky-noise removal, and other filtering during the reduction). The $50''$ FWHM effective PSF bowl is the direct result of explicit large-scale-structure

(LSS) filtering during the reduction with a 50" FWHM Gaussian profile. We checked the Gaussian main beam assumption on quasars (detected at high signal-to-noise >1000), and confirmed that >98% of the integrated flux inside a $R=50''$ circular aperture is recovered in the 17.5" FWHM beam-smoothed peak during night-time observations (i.e. for all of our data). We also checked that the filtered PSF accurately recovers the quasar fluxes, when these were LSS filtered the same as our deep field map, and found no further degradation of photometric accuracy associated with the LSS filtering. Therefore, we are very confident that we have sufficient understanding of the effective PSF in our map, and that the systematic errors of the extracted points source fluxes are kept below 2%.

The source extraction code we used is part of the CRUSH software package, and is the same source extraction tool that was used and described in Weiß et al. (2009). It implements an iterated false-detection rate algorithm. Apart from peak position and flux, the algorithm calculates an estimated confidence and an expected cumulative false detection rate for each extracted source. We caution that the confidence levels and false-detection rates are guiding values only, which represent our best statistical estimate without prior knowledge of the 2-mm source counts. A more accurate characterization of confidence levels and/or cumulative false-detection rates would require accurate prior information of the true counts of the 2-mm source population.

4.1.1. Overview of the CRUSH source extraction tool.

Here, we offer a concise summary of the approach implemented by CRUSH 'detect' tool, which we used for the source extraction.

The expected false detection rate, i.e. the expected number of pure noise peaks mistakenly identified as a source, is given by $N_f(\Sigma) = NQ(\Sigma)$, where N is the number of independent Gaussian variables in the map, and $Q(\Sigma) = 1 - P(\Sigma)$ is the complement cumulative Gaussian probability, i.e. the probability of measuring a deviation larger than a chosen significance, Σ . A smoothed and filtered map with extraction area A contains

$$N \approx \frac{4.5A}{1.13 (\Delta_{\text{smooth}}^2)} \left(1 - \frac{\Delta_{\text{smooth}}^2}{\Delta_{\text{filter}}^2} \right) \quad (2)$$

independent variables in terms of the FWHM widths of the Gaussian smoothing ($\Delta_{\text{smooth}} = \Delta_{\text{beam}}$) and the applied large-scale filtering of the map ($\Delta_{\text{filter}} = 50''$). The right-hand-side term in the formula accounts for the lost degrees of freedom due to the explicit spatial filtering of our map. The approximately 4.5 parameters per smoothing beam were determined empirically based on the occurrence of significant noise peaks in simulated noise maps. The formula was verified to yield close to the expected number of false detections in simulated noise maps with varying areas and filtering properties, and with N_f targeted between 0 to 1000. Thus, the above expression will accurately predict the actual false detection rate, as a function of detection threshold, as long as the map noise is known precisely. For our map, with ~ 321 GISMO beams, a 2.99σ cut yields $N_f(2.99) \approx 2$ expected false detections.

Due to the presence of many resolved but undetected sources in the map (asymmetric confusion), our noise estimates are bound to be slightly overestimated (even with the median-noise based estimate used). To our best knowledge, all statistical estimates of map noise, which are based on the observed map itself, will result in overstated noise estimates in the presence of asymmetric confusion (resolved sources below detection). Neither the jackknifed noise, nor radiometric noise, can help offer better estimates, as the extraction noise should include the effect of symmetric confusion (unresolved faint sources) beyond what these can offer. As a result of an inevitably biased noise estimation process, the corresponding false detection rate estimates are slightly above actual, and represent a useful conservative upper bound. This is confirmed by the simulations, presented in section 4.5, which found that if the 2 mm source counts were those of, e.g., Béthermin et al. (2011) or Lapi et al. (2011), then the actual false detection rate would be 1.34 or 0.55, respectively, vs. the expected 2. However, as we stated earlier, we cannot unbiased our noise estimates, or quantify the true false detection rate, without prior knowledge of the true 2-mm source counts, which are not well-constrained at present. Instead, our estimates offer strong upper bounds for the unknown actual false-detection rates.

Each source identified above the significance cut is removed from the map with the smoothed and filtered point-spread function before the extraction proceeds. Subtraction with the filtered PSF allows the detection of further nearby peaks, which may have been previously suppressed by the negative filter bowl surrounding the previous detections. The circular area ($r^2 = \Delta_{\text{beam}}^2 + \Delta_{\text{smooth}}^2$) containing the main beam of the detected source is flagged after the extraction, since it no longer contains meaningful information after the removal of the source from within. To ensure that our catalog is based on the most accurate measure of the map noise and zero levels, CRUSH estimates the zero level using the mode of the map flux distribution, and estimates the noise from the median observed deviation $\text{median}(x^2) \approx 0.454937\sigma^2$. Both measures are relatively robust and reasonably unbiased by the presence of relatively bright sources, or localized features, in the map.

For each source candidate, CRUSH estimates a detection confidence based on the expected false detection rate N_f . According to Poisson statistics, the detection confidence C of a single peak is the probability that no such peak occurs randomly, i.e. $P_0(N_f) = e^{-N_f}$. This is then further refined to include information from other sources already detected in the map. Thus, if n true sources with apparent significance above Σ are known *a priori* to exist in the map, than any given peak at significance Σ may be one of n sources, or one of the N_f expected false detections, hence the probability of false detection for each of $n + N_f$ peaks is reduced by a factor of $N_f/(n + N_f)$. (In other words, we should expect only N_f false detections (noise peaks) for every n actual sources detectable above a given threshold.) CRUSH uses the number of sources $N(>\Sigma)$ that were already extracted above significance Σ minus the expected false detection rate $N_f(\Sigma)$ as a self-consistent proxy for n , which is a reasonable assumption when prior knowledge of the actual underlying counts is

not readily available (as in our case). As such, the individual confidence levels of consecutive detections are estimated as

$$C(\Sigma) = 1 - \frac{N_f(\Sigma)}{N(>\Sigma)} \left(1 - e^{-N_f(\Sigma)}\right). \quad (3)$$

4.2. Deboosting

Deboosting is a statistical correction to the observed flux densities, when source counts fall steeply with increasing brightness (e.g. Crawford et al. (2010) and references therein). Thus, in a statistical ensemble of sources, the same observed flux arises more often from one of many fainter sources than from the few brighter ones, relative to the measured value. We assume a measurement with Gaussian noise (validated by the closely Gaussian jackknife noise distribution) and a 2 mm source count model scaled from observationally constrained 850 μm counts (e.g. Coppin et al. (2006), Weiß et al. (2009)) assuming $T_d/(1+z) \gtrsim 10$ K (Kovács et al. 2006a) and dust emissivity index (β) of 1.5 (Kovács et al. 2010). We also deboosted our data using the physical number-count models of Lapi et al. (2011) and Béthermin et al. (2011), see section 5.1.

For deboosting we followed a Bayesian recipe, such as described in Coppin et al. (2005, 2006):

$$p(S_i|S_o, \sigma) \propto p(S_i) p(S_o, \sigma|S_i) \quad (4)$$

expressing the probability of intrinsic source flux S_i in terms of the observed flux S_o and its measurement uncertainty σ .

However, we made some important modifications to the recipe to account for the possibility that the observed flux arises from multiple overlapping galaxies, and we account for confusion. Accordingly, we replace the single isolated source assumption $p(S_i) \propto \frac{dN}{dS}(S_i)$ of Coppin et al. (2005, 2006), with the compound probability that one or more (up to m) resolved sources in the beam contribute to an aggregated intrinsic flux S_i :

$$p(S_i) = \int_0^{S_i} dS_1 \dots \int_0^{S_{m-1}} dS_m \pi(S_1) \dots \pi(S_m) \delta\left(S_i - \sum_{k=1}^m S_k\right). \quad (5)$$

Inside the integrals is the product of the individual component probability densities $\pi(S_k)$, which correspond to S_i arising from a specific combination of $(S_1 \dots S_m)$ individual components. The delta function ensures that the component fluxes considered add up to the total intrinsic flux S_i when integrated. Each nested integral for $S(k)$ is performed up to the previous flux S_{k-1} , indicating that each successive component S_k is no brighter than the previous one S_{k-1} , and ensuring that each particular combination of fluxes is counted one time only. Once the fluxes in the outer integrals sum up to S_i , the remaining inner integrals can be skipped (in numerical implementations) corresponding to fewer than m actual contributors (or to keep to a more formal notation we can add the delta function at zero, i.e. $\delta(S_k)$, to the definition of $\pi(S_k)$ below to achieve exactly without omitting the any inner integrals). When overlaps are ignored ($m=1$), Eq. 5 reduces to $p(S_i) = \pi(S_i)$ naturally (no integration required).

The differential source counts $\frac{dN}{dS}(S_k)$ determine the probability that there is at least one intrinsic source with brightness S_k in the beam, resolved or unresolved. The distinction between resolved and unresolved sources is important: unresolved sources cause a symmetric widening of the map noise distribution (confusion noise) compared to the experimental noise (e.g. radiometric down-integration as measured by a jackknife); resolved sources, on the other hand, detected or not, will manifest as an excess of flux on the positive side of the observed flux (or S/N) distribution. Deboosting naturally needs to consider resolved sources only.

If one distributes N sources randomly in some large area A (for simplicity's sake let's consider the same area to which the counts are normalized, whether deg^2 or sr) with $N_b = A/2\pi\sigma_b^2$ ($N_b \gg 1$) independent beams (FWHM $\approx 2.35\sigma_b$), then the chance that none of these sources fall into a given beam (our detection beam) is:

$$q(N) = \left(1 - \frac{1}{N_b}\right)^N \approx e^{-N/N_b} \quad (6)$$

Therefore, at a given flux density S_k the probability density of a source with brightness S_k being resolved (i.e. unblended with brighter ones) is

$$\pi_+(S_k) \approx \frac{1}{N_b} \frac{dN}{dS}(S_k) e^{-N(>S_k)/N_b}, \quad (7)$$

in terms of the integrated number counts $N(>S)$, and the corresponding differential counts $\frac{dN}{dS}$. Here, $\pi_+(S_k)$ measures the probability density that at least one resolved source with flux S_k falls inside the detection beam, and does not exclude the possibility of further fainter components within the same beam (hence the plus sign as the subscript). Using $\pi_+(S_k)$, however, we can easily express the probability density $\pi(S_k)$ for exactly one component with S_k in a given beam, by simply subtracting the integrated probability that there is at least one other fainter object in that same beam with the first one:

$$\pi(S_k) = \pi_+(S_k) \left(1 - \int_0^{S_k} \pi_+(S) dS\right). \quad (8)$$

For the highest order m under consideration, we may truncate by setting $\pi(S_m) \approx \pi_+(S_m)$. The approximation is valid as long as m is chosen to be large enough such that resolved overlaps with further components (the right-hand integral term) are negligible.

For the particular case of the GISMO 2 mm sources, we considered up to 3 overlapping components ($m=3$) contributing to the observed fluxes. We verified that this was sufficient, as we noticed no measurable degree of incremental change in the deboosted values (and profiles) between $m=2$ and $m=3$. We chose $m=3$ to be on the safe side. At the same time allowing for at least 2 overlapping components instead of just a single isolated source ($m=1$) did have a significant impact on the deboosting results, justifying our modified approach.

Since our source extraction algorithm determines the map zero level as the mode (not the mean) of the distribution, the extracted source fluxes are easily measured

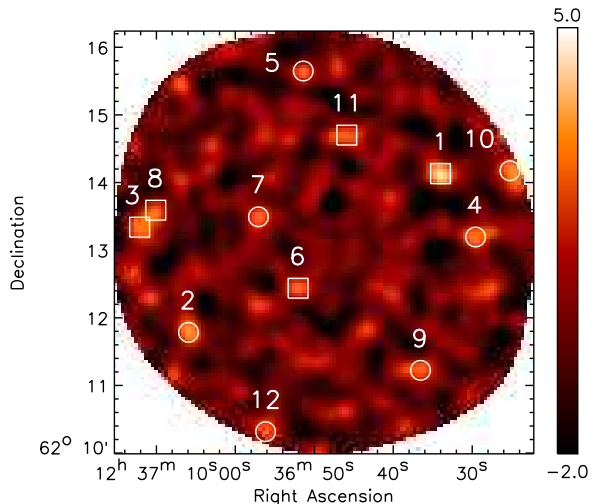


FIG. 5.— Map of the S/N for the smoothed and filtered GISMO Deep Field (GDF). The applied point source filter yields an effective image resolution of $24.7''$ FWHM. Sources at $S/N > 3$ (Table 1) are marked with circles (sized as the $17.5''$ FWHM of the diffraction limited GISMO beam) and squares (for sources with 1.2 mm and/or 850 μm counterparts).

against the unresolved background. And, because our deboosting method is based on resolved sources only, it also means that no additional zero-level adjustment is necessary. As a result, the distribution naturally does not extend to negative fluxes, as is demonstrated by the posterior probability distributions of the extracted GISMO sources shown in the appendix.

4.3. Extraction – Results

Figure 5 shows the beam-smoothed signal-to-noise map of the 2mm GDF. Figure 6 shows the noise map, demonstrating that the innermost $4'$ of the GDF observations have an rms of between 130 and 140 μJy . The source extraction was performed out to twice this level, i.e. up to $\sim 260\mu\text{Jy}$ rms (Fig. 6). The area for the source extraction is ~ 31 square arcminutes (~ 321 GISMO beams).

The source extraction algorithm finds 12 positive sources and 3 negative “sources”. The number of significant negatives is consistent with the expected false detection rate of $N_f = 2$ for a 2.99σ detection threshold in the extraction area (see Section 4.1). The measured 2 mm fluxes range from 400 to 870 μJy .

Table 1 shows the measured fluxes of the 12 positive sources with the achieved signal-to-noise, the estimated detection confidence levels, and the expected cumulative false detection rate N_f . The fluxes presented in the table show the measured fluxes, while Table 2 shows the de-boosted values, using a scaled version of the broken power-law SHADES galaxy number counts (Coppin et al. 2006), the Lapi et al. (2011) number counts, as well as the values corresponding to the Béthermin et al. (2011) counts.

In order to identify (sub)millimeter counterparts of our sources we cross correlated our data with the 1.16 mm combined MAMBO/AzTEC source catalog containing the 1.1 mm the AzTEC data from Perera et al. (2008), and the MAMBO source catalog from Greve et al. (2008), as well as the SCUBA

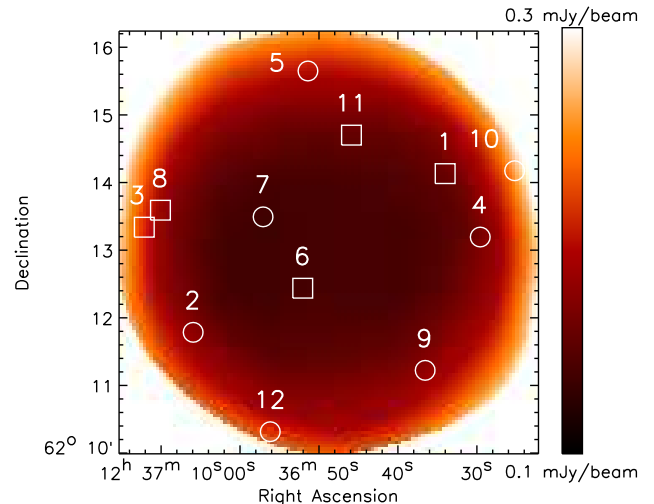


FIG. 6.— Noise map for the smoothed and filtered GISMO Deep Field (GDF). The noise in the innermost $\sim 4'$ diameter in the map is $\sim 135\mu\text{Jy}/\text{beam}$. Sources at $S/N > 3$ (Table 1) are marked with circles (sized as the $17.5''$ FWHM of the diffraction limited GISMO beam) and squares (for sources with 1.2mm and/or 850 μm counterparts).

850 μm source catalog of GOODS-N (Borys et al. (2003), as well as Borys et al. (2004), Pope et al. (2005), and Pope et al. (2006)). Table 3 shows the measured and deboosted 1.2 mm and 850 μm fluxes of the sources with counterparts, if available. The equation we used for the maximum allowable separation between the GISMO/MAMBO/AzTEC/SCUBA sources in order to be considered a counterpart is given by:

$$r_{max}^2 = 4\sigma_p^2 - 2\sigma_{beam}^2 \ln \left(1 - \frac{2}{\text{SNR}} \right), \quad (9)$$

where r_{max} is the search radius where the counterpart must fall with $> 98\%$ confidence, σ_p is the 1 sigma catalog position error combined with the rms astrometric accuracy of our map ($< 0.8''$) in quadrature, σ_{beam} is the 1 sigma beam size $\sim FWHM/2.35$, and SNR is the observed signal-to-noise ratio of the detection. When searching for a counterpart in another low S/N dataset of (sub-)mm data set, a second identical SNR-dependent term needs to be added to the search-radius expression above, reflecting the inherent positional uncertainty of the other known mm-wave detection. In section 4.5, we demonstrate the applicability of equation 9 for our dataset.

Combining the GISMO, MAMBO/AzTEC/SCUBA observations (Fig. 7), we identify three additional sources with detection confidence level of $> 80\%$. These are tabulated in Table 4 and 5. The median and mean redshifts of all sources with counterparts and known redshifts are $\bar{z} = 2.91 \pm 0.94$ and $\bar{z} = 3.3$, respectively.

4.4. GDF sources without a counterpart

Seven of the detected 2mm sources have no counterpart in either the MAMBO/AzTEC or the SCUBA data (GDF-2000.2, GDF-2000.4, GDF-2000.5, GDF-2000.7, GDF-2000.9, GDF-2000.10, and GDF-2000.12). Considering their observed signal to noise ratio and the three negative detections in the map, plus the statistical expectation of 2 false detections as derived in section 4.1, the

TABLE 1
GDF 2MM SOURCES

ID	RA [J2000]	DEC [J2000]	S [mJy]	S/N	confidence [%]	N_f
GDF-2000.1	12:36:33.98	62:14:08.0	0.79 ± 0.14	5.53	100	0.000
GDF-2000.2	12:37:05.95	62:11:47.2	0.67 ± 0.16	4.18	99	0.021
GDF-2000.3	12:37:12.17	62:13:20.4	0.78 ± 0.19	4.10	99	0.029
GDF-2000.4	12:36:29.54	62:13:11.7	0.53 ± 0.15	3.51	87	0.330
GDF-2000.5	12:36:51.40	62:15:39.1	0.61 ± 0.18	3.37	87	0.541
GDF-2000.6	12:36:52.06	62:12:26.4	0.42 ± 0.13	3.33	89	0.627
GDF-2000.7	12:36:57.09	62:13:29.7	0.40 ± 0.12	3.24	89	0.875
GDF-2000.8	12:37:10.12	62:13:35.5	0.54 ± 0.17	3.22	89	0.918
GDF-2000.9	12:36:36.54	62:11:13.5	0.51 ± 0.16	3.17	84	1.117
GDF-2000.10	12:36:25.16	62:14:10.5	0.87 ± 0.29	3.06	84	1.630
GDF-2000.11	12:36:45.88	62:14:42.2	0.43 ± 0.14	3.04	84	1.735
GDF-2000.12	12:36:56.17	62:10:19.1	0.54 ± 0.19	3.02	84	1.828

NOTE. — N_f is the expected cumulative false detection rate as defined in Section 4.1. The maps are beam-smoothed (by 17.5'' FWHM) to an effective 24.7'' FWHM image resolution for point source extraction

TABLE 2
DEBOOSTED 2MM FLUX DENSITIES ASSUMING DIFFERENT
SOURCE COUNTS OR MODELS

ID	deboosted S' [mJy]		
	Lapi et al.	Coppin et al.	Bethermin et al.
GDF-2000.1	0.75 ± 0.16	0.69 ± 0.14	0.71 ± 0.16
GDF-2000.2	0.56 ± 0.21	0.53 ± 0.18	0.51 ± 0.20
GDF-2000.3	0.63 ± 0.25	0.59 ± 0.22	0.56 ± 0.25
GDF-2000.4	0.39 ± 0.20	0.37 ± 0.18	0.35 ± 0.19
GDF-2000.5	0.40 ± 0.24	0.38 ± 0.21	0.34 ± 0.23
GDF-2000.6	0.31 ± 0.16	0.29 ± 0.15	0.28 ± 0.16
GDF-2000.7	0.29 ± 0.15	0.27 ± 0.14	0.26 ± 0.15
GDF-2000.8	0.35 ± 0.22	0.33 ± 0.20	0.30 ± 0.20
GDF-2000.9	0.33 ± 0.20	0.31 ± 0.19	0.28 ± 0.19
GDF-2000.10	0.40 ± 0.33	0.37 ± 0.30	0.30 ± 0.29
GDF-2000.11	0.28 ± 0.17	0.26 ± 0.16	0.25 ± 0.17
GDF-2000.12	0.33 ± 0.23	0.31 ± 0.21	0.24 ± 0.20

NOTE. — Derived by extrapolation of the Coppin et al. (2006) SHADES number counts to 2 mm counts, or from the Lapi et al. (2011) or Béthermin et al. (2011) models.

TABLE 3
GDF 2 MM SOURCES WITH (SUB-)MM COUNTERPARTS

ID	z	ID	$S(850\mu\text{m})$ [mJy]	$S'(850\mu\text{m})$ [mJy]	ID	$S(1.2\text{ mm})$ [mJy]	$S'(1.2\text{ mm})$ [mJy]	δr [']
GDF-2000.1	4.042	GN 850.10	11.3 ± 1.6	8.6 ± 4.8	AzGN03	5.2 ± 0.6	4.7 ± 1.7	6.0
GDF-2000.3	1.992	GN 850.39	7.4 ± 2.0	3.8 ± 2.8	AzGN07A*	blend*	...	9.8
					GN1200.3	3.9 ± 0.6	3.3 ± 1.3	
GDF-2000.6	5.183	GN 850.14	5.9 ± 0.3	5.9 ± 0.3	AzGN14	2.9 ± 0.6	2.2 ± 1.0	1.6
GDF-2000.8	1.97		< 6	...	AzGN07B*	blend*	...	13.8
GDF-2000.11	2.30	GN 850.12	8.6 ± 1.4	6.4 ± 3.6	AzGN08	3.0 ± 0.6	2.4 ± 1.1	6.6

NOTE. — MAMBO/AzTEC 1.2 mm data are from Greve et al. (2008) and Perera et al. (2008), SCUBA 850 μm data from Borys et al. (2003), Pope et al. (2005), and Greve et al. (2008). z is the measured redshift. The deboosted flux values S' are based on the SHADES counts, using the same equations used for calculating the 2 mm data counts in order to be consistent with the deboosting of the 2 mm fluxes shown in Table 2.

* Both, GDF-2000.3 and GDF-2000.8, are associated with AzGN07, which implies that this 1.2 mm source is a blend of two sources.

TABLE 4
LOW S/N GDF 2MM SOURCES – IDENTIFIED THROUGH (SUB-)MM COUNTERPARTS

ID	RA [J2000]	DEC [J2000]	$S(2\text{mm})$ [mJy]	z	ID	λ	S [mJy]	S' [mJy]	δr [']
GDF-2000.13	12:37:12.01	62:12:14.0	0.52 ± 0.20	2.91	GN 850.21	850 μm	5.7 ± 1.2	3.8 ± 2.3	8.1
GDF-2000.14	12:36:27.40	62:12:13.8	0.50 ± 0.19	4.69	GN 1200.29	1200 mm	2.6 ± 0.6	1.9 ± 0.9	0.0
GDF-2000.15	12:36:45.00	62:11:47.1	0.36 ± 0.15	...	AzGN10	1200 mm	2.6 ± 0.6	1.9 ± 0.9	4.9
					GN 850.28	850 μm	1.7 ± 0.4	0.6 ± 0.6	1.8

NOTE. — z is the measured redshift. Deboosted flux values S' were calculated directly (850 μm) from or extrapolated (1.2 mm) from the SHADES number counts.

TABLE 5
DEBOOSTED 2MM FLUX DENSITIES ASSUMING DIFFERENT
SOURCE COUNTS OR MODELS

ID	deboosted S' [mJy]		
	Coppin et al.	Lapi et al.	Bethermin et al.
GDF-2000.13	0.26 ± 0.21	0.24 ± 0.20	0.21 ± 0.19
GDF-2000.14	0.25 ± 0.20	0.23 ± 0.19	0.20 ± 0.18
GDF-2000.15	0.21 ± 0.15	0.17 ± 0.14	0.17 ± 0.14

NOTE. — Derived by extrapolation of the Coppin et al. (2006) SHADES number counts to 2 mm counts, and from the Lapi et al. (2011) or Béthermin et al. (2011) models.

TABLE 6
GDF 2MM SOURCES – DERIVED PARAMETERS

ID	$ \chi^2 $	T_c [K]	$\log M_d$ [M_\odot]	$\log L$ [L_\odot]	q_L	q_{IR}	τ_{peak}
GDF 2000.1	0.70	51.2 ± 2.0	8.53 ± 0.07	13.52 ± 0.06	2.83 ± 0.21	3.08 ± 0.21	1.28
GDF 2000.3	1.08	40.8 ± 0.7	8.59 ± 0.14	13.10 ± 0.03	2.68 ± 0.10	2.91 ± 0.10	1.08

NOTE. — All quantities were fitted using the measured redshifts (Table 2), temperature-distribution model ($dM_d/dT \propto T^{-7.2}$) with $\beta = 1.5$ and assuming a 2 kpc emission diameter. The dust masses assume $\kappa(850\mu\text{m}) = 0.15 \text{ m}^2 \text{ kg}^{-1}$. Uncertainties are 1σ total errors of the fits to data, which do not include the uncertainties in the redshift values. The following quantities are shown in the table: $|\chi^2|$ residual scatter around the fit, T_c temperature of the dominant cold component, M_d dust mass, L integrated IR luminosity, q_L radio-(F)IR correlation constant as defined in Kovács et al. (2010), q_{IR} radio-(F)IR correlation constant as defined in Ivison et al. (2010), and τ_{peak} optical depth around the IR emission peak.

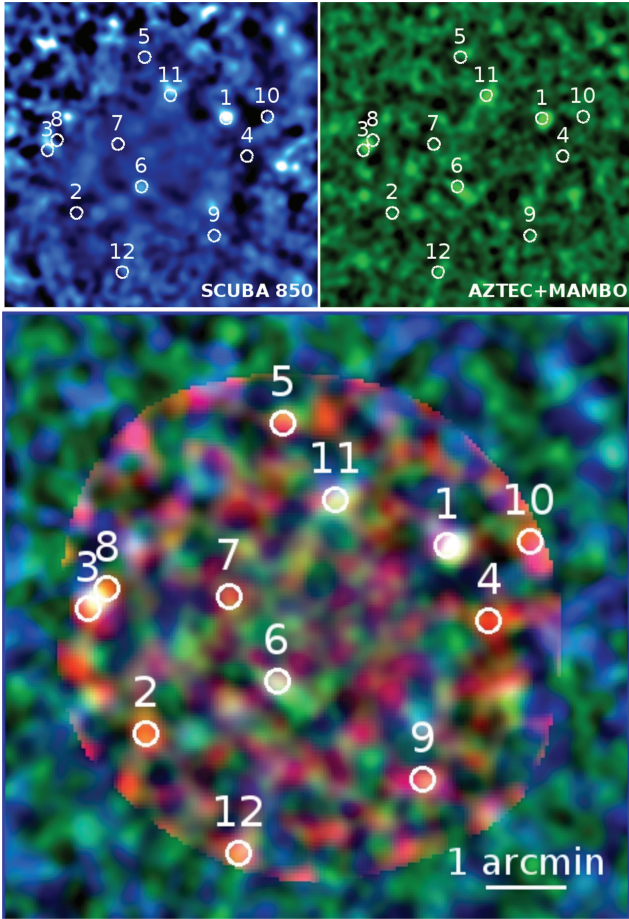


FIG. 7.— The GISMO Deep Field Sources at multiple wavelengths. The SCUBA 850 μm map (Pope et al. 2005) is shown in blue. The AzTEC+MAMBO 1.2 mm map (Penner et al. 2011), is shown in green. The 2 mm GISMO image is shown in red.

most likely scenario is that 5 of these are real detections.

Figure 8 shows the redshift-dependent GDF equivalent map sensitivities at 850 μm and 1.2 mm for the detection of a galaxy with the SED shown in Fig. 1. The figure demonstrates that the equivalent 1.2 mm source sensitivity requirement to match the GDF source sensitivity varies by about a factor of two with redshift, whereas the required depth of the 850 μm SCUBA map varies significantly. In order to achieve the same GDF source sensitivity the 1 σ map noise rms at 1.2 mm would need to be ~ 0.7 mJy/beam if the galaxy were at a redshift of 2, while it would require ~ 0.4 mJy/beam if the same source were at an extremely high redshift. At 850 μm the matching map flux sensitivity requirement ranges from about 2 mJy/beam for a source at a redshift $z \sim 2$ to ~ 0.5 mJy/beam for a source at extremely high redshifts. Table 7 shows the actual 850 μm and 1.2 mm map sensitivities at the seven positions of GDF sources without counterparts. The table shows that the depth of the 1.2 mm map is quite homogeneous for all sources without counterparts with an rms of about 0.56 mJy for each of those. This means that for our template SED, the source sensitivity of the GDF map exceeds that of the 1.2 mm data for redshifts of about $z = 6$ and greater. The situation for the SCUBA data is different, since the sensitivity over that map varies very significantly as is demonstrated by the range as shown in the same table.

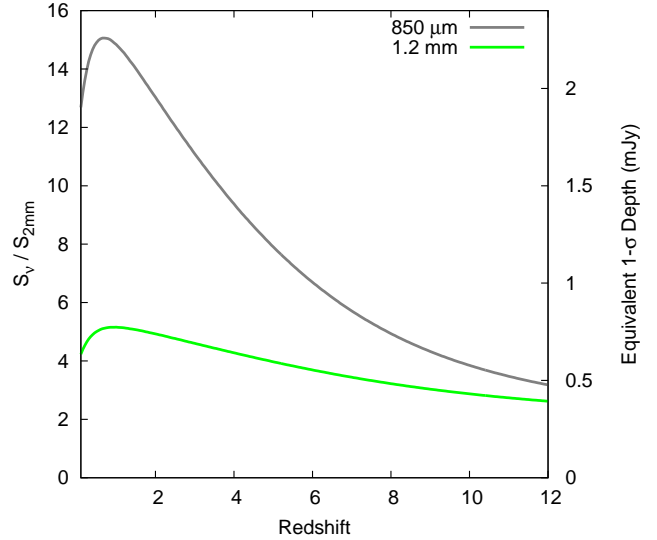


FIG. 8.— Redshift-dependent GDF colors at 850 μm and 1.2 mm, relative to the 2 mm fluxes, based on the same SED model as Fig. 1. On the right axis, we indicate the typical 1- σ depths that an 850 μm or 1.2 mm map would have to reach, as a function of redshift, to provide equivalent coverage of the sources to that of the GISMO deep field in this paper.

Only at the position of GDF-2000.7 the sensitivity of the SCUBA map is below an rms of 1 mJy/beam (Table 7). For the other sources the 850 μm map sensitivities are between 1 and 2 mJy/beam rms. A comparison of these numbers with the equivalent 850 μm GDF point source sensitivities shown in Fig. 8 shows that the 2 mm point source sensitivity of all GDF sources without counterparts, with the exception of GDF-2000.7, exceeds the point source sensitivity of the SCUBA map for redshifts $z > 6$. Taken together, the 850 μm and 1.2 mm data on non-detected GDF sources are entirely consistent with sources at $z > 6$, assuming the template from Fig. 1 applies. However, another aspect when considering that a detected 2 mm source has no counterpart in another catalog is that of completeness or the probability of a low S/N source to be extracted. For the 850 μm SCUBA map, e.g., only the completeness of sources with $S/N \gtrsim 5$ is $> 90\%$ (Coppin et al. 2006), i.e. the S/N dependent probabilities are similar to those derived in the completeness analysis for the GDF data presented in section 4.5. The situation is similar for the 1.2 mm data. As a result, the S/N limitations of the 850 μm SCUBA and 1.2 mm MAMBO/AzTEC data sets combined with the low number of 2 mm sources prevents us from deriving significant redshift constraints on the GISMO sources without counterparts. Larger deep 2-mm surveys will be needed to photometrically study what fraction of the 2-mm population is at very high-redshift ($z > 6$), one of the main science goals of GISMO.

4.5. Simulations

Additional characterizations of the extracted sources were obtained through simulations, generally following the analysis of Weiß et al. (2009). These simulations demonstrate also the validity of the equations used for source extraction presented in sections 4.1 and 4.3. The simulations started with the construction of 100 variations of jackknifed noise maps that were generated from the original data. These maps provide

TABLE 7
850 μm AND 1.2 MM SENSITIVITIES AT THE
POSITIONS OF THE GDF 2 MM SOURCES
WITHOUT COUNTERPARTS

ID	$rms(850\mu\text{m})$ [mJy/beam]	$rms(1.2\text{ mm})$ [mJy/beam]
GDF-2000.2	1.02	0.56
GDF-2000.4	1.54	0.57
GDF-2000.5	1.64	0.55
GDF-2000.7	0.42	0.55
GDF-2000.9	1.09	0.56
GDF-2000.10	2.04	0.56
GDF-2000.12	1.85	0.55

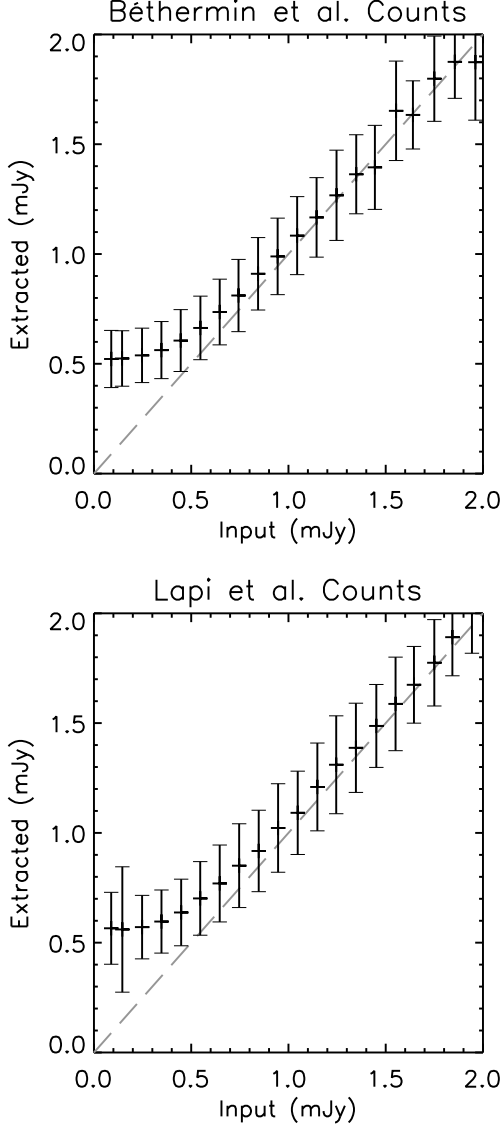


FIG. 9.— Plots of the extracted source flux density as a function of the input flux density for simulations using Béthermin et al. (2011) source counts (top) and extrapolated Lapi et al. (2011) source counts (bottom). The error bars reflect the 1σ dispersion for flux density intervals populated by more than 1 simulated source. The boosting of the flux densities is evident for sources at < 1 mJy.

accurate representations of the noise of the observations. We next constructed 3000 versions of the point source distribution across the full area of the map (not just

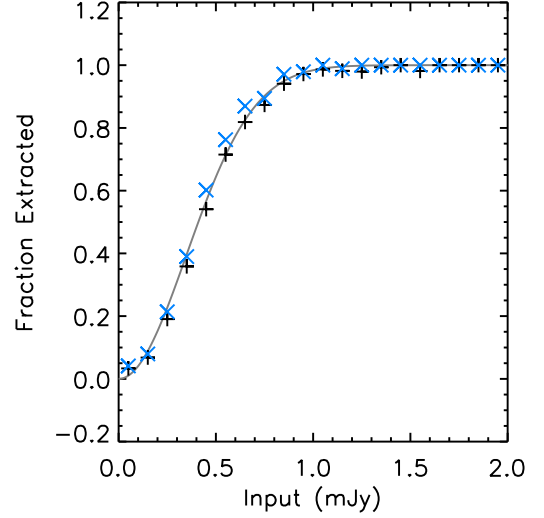


FIG. 10.— The completeness (fraction of simulated input sources that are recovered by the source extraction) is plotted as a function of the input flux density of the simulated sources. Simulations using Béthermin et al. (2011) number counts and extrapolated Lapi et al. (2011) counts are shown as blue and black respectively. In both cases the completeness drops to 50% at 0.4 mJy. The completeness can be empirically fit by the inverted gaussian function $1 - \exp(-0.5(S/\sigma)^2)$ where $\sigma = 0.348$ mJy (gray line).

the low noise, high coverage regions) based on the Béthermin et al. (2011) model and the smoothed and filtered GISMO beam. The simulations included sources down to 0.01 mJy in order to provide an appropriate confusion background. Source brightnesses were chosen at random from the cumulative source counts, $N(> S)$, producing a sample size appropriate for 3000 images. These sources were then distributed randomly across the full set of 3000 noise-free images. Therefore, the peak brightness and the total number of sources in each simulated image are subject to variation due to Poisson statistics. There were an average of 2264 sources in each simulated image. The jackknife noise maps were added to the simulations, reusing each of the noise maps 30 times. The 3000 simulated images were then run through the source detection procedure, using the same settings as were applied to the actual data. These procedures were repeated using another set of 3000 simulated point source maps based on the Lapi et al. (2011) models, which predict a higher source density with a mean of 3780 sources ($S > 0.01$ mJy) per simulated image.

The positions of the extracted sources were matched

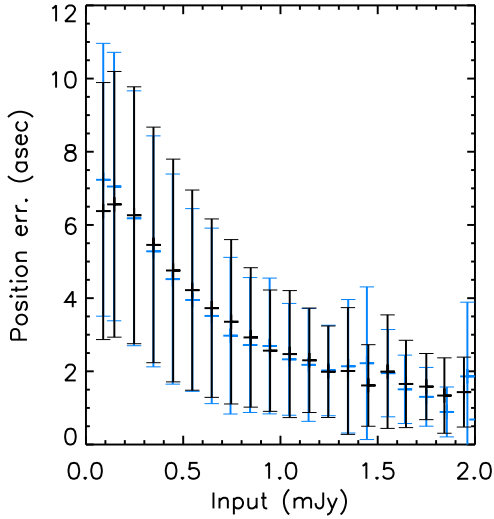


FIG. 11.— The mean positional errors with $1 - \sigma$ error bars of extracted sources are plotted as a function of the input source brightness. The simulations use the Béthermin et al. (2011) number count model and the extrapolated Lapi et al. (2011) model, shown as blue and black respectively. The mean positional errors are $< 5.1''$ for sources brighter than 0.4 mJy.

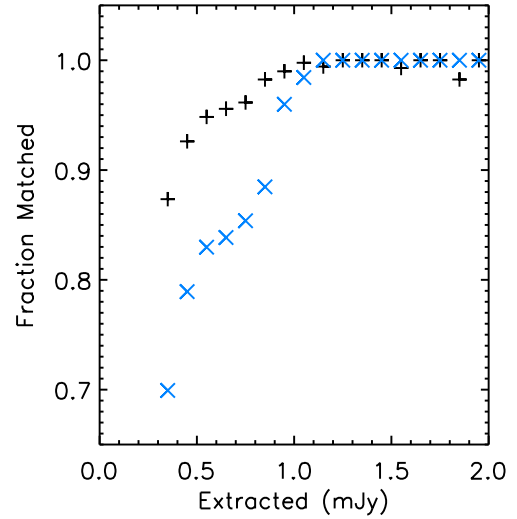


FIG. 13.— The reliability (fraction of extracted sources that match input simulated sources) is plotted as a function of the input flux density of the simulated sources. The simulations use the Béthermin et al. (2011) number count model and the extrapolated Lapi et al. (2011) model, shown as blue and black respectively.

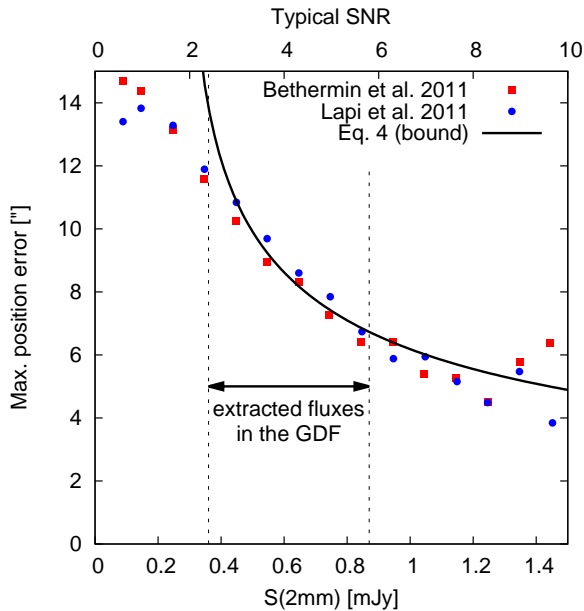


FIG. 12.— Determining search radii: $2 - \sigma$ position errors of sources extracted from the simulations as a function of input flux (points), vs the prediction from Eq. 9 as a function of S/N (solid curve). To convert the simulated input fluxes to S/N, we estimated an average depth of 0.15 mJy/beam in our map, and we used $\sigma_p = 0$ since there were no intrinsic pointing errors in the simulated data. The dotted lines indicate the range of fluxes extracted from our map, i.e. the range for which search radii are calculated. We note that at low S/N the simulations tend to falsely identify a nearer chance peak (one of the many faint sources filling the map, or a noise peak) as the counterpart to the input. As such, the simulations tend to underestimate the true position errors at low S/N, explaining the deviation from the curve below detection level. Also, the asymptotic behavior at high S/N (> 10) is not well modeled by Eq. 9, however for the range of fluxes considered, the calculated search radii are in remarkable agreement with simulations, thus justifying our reliance on them.

with those of the simulated input sources for each of the 3000 simulations based on the two different models. For a given extracted source, the matching input simulated source was chosen as the brightest source within a 5 pixel $= 15''$ radius and with a brightness > 0.1 mJy. In most cases there is only one source within $15''$, but infrequently a faint input source happens to lie closer to the extracted source position than a brighter input source that is the true origin of the extracted source.

Figure 9 compares the brightnesses of the extracted sources with the associated input source brightnesses. Results are binned in 0.1 mJy intervals, and error bars indicate the standard deviation of the sources averaged in each bin. For either set of simulations, we find that source boosting becomes significant at $S \lesssim 1$ mJy. Consistent with the fluxes we measure for the extracted sources from the GDF, typical fluxes extracted from the simulations are greater than ~ 400 uJy/beam, as we have used the same detection criterion (at $2.99 - \sigma$) on our simulated data as on our observed map, and because our map has an average $1 - \sigma$ depth of around 135–140 mJy/beam. Figure 10 shows the completeness of the extracted sources, or the fraction of the input sources that are extracted. Whereas essentially all sources with $S > 1$ mJy are recovered by the source extraction, the completeness drops to $\sim 50\%$ at $S \sim 0.4$ mJy for both models of the source counts. Figure 11 shows the expected trend of increasing positional errors with decreasing source brightness. The positions of fainter sources are more strongly affected by noise and source confusion, consistent with Eq. 9. In Figure 12 we compare the search radius given by Eq. 9 with the results from the simulated positional accuracy. The figure demonstrates that over the relevant range of fluxes, the calculated search radii are in remarkable agreement with simulations, thus justifying our reliance on them. An assessment of the reliability of the source extractions is shown in Figure 13, where we plot the fraction of sources

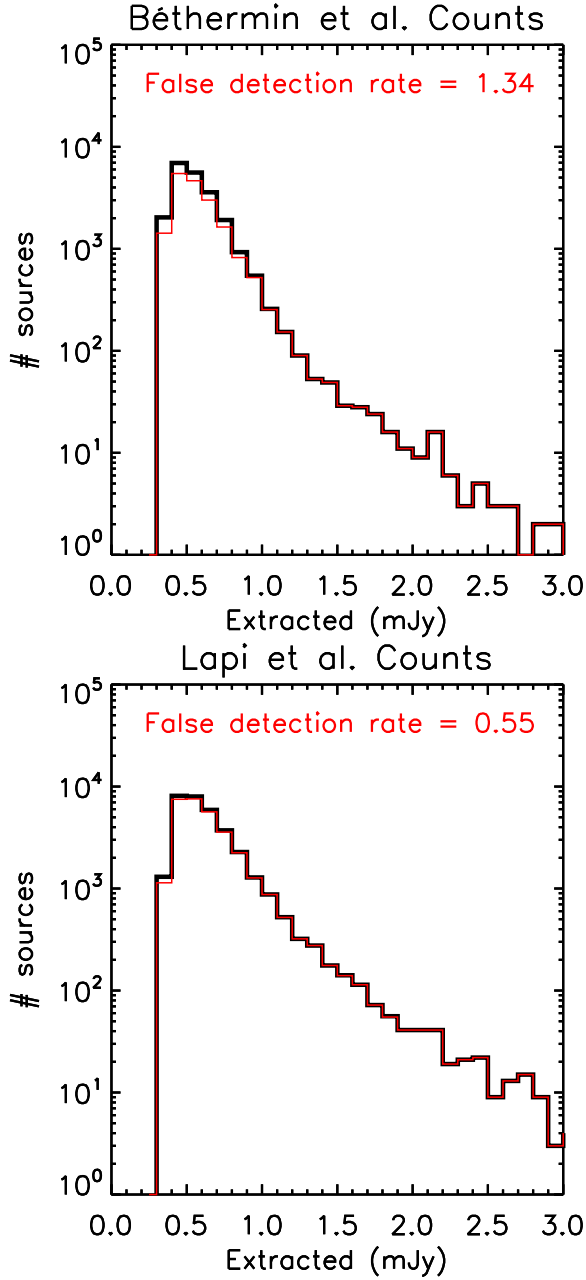


FIG. 14.— Histograms of the brightnesses of the extracted sources (black) and the simulated sources (red). False sources are indicated by an excess number of extracted sources. The mean number of false sources extracted is ~ 1 , but the value is inversely related to the number density of the simulated sources.

at a given extracted brightness that can be associated with corresponding input source in the simulations. For either model, the reliability of the extracted sources begins to drop at $S < 1.2$ mJy. The drop is somewhat greater for simulations using the Béthermin et al. (2011) source counts, which has an overall lower density of sources. As expected, the reliability drops more rapidly near the detection threshold (~ 400 mJy/beam, given the average $1-\sigma$ depth around 135–140 mJy/beam over the full map). Note, that because the reliability depends strongly on the shape of the unknown actual number counts, especially near the chosen detection threshold, the confidence levels shown in Table 1, and calculated

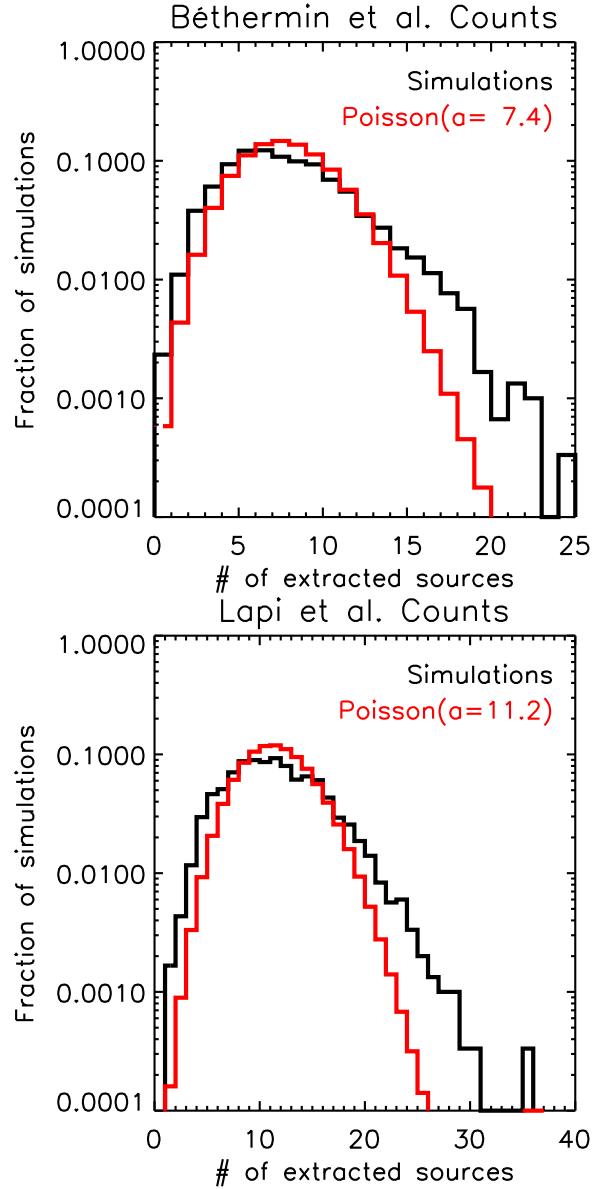


FIG. 15.— Histograms of the number of sources extracted per field, over the set of 3000 simulated fields. Overplotted is the Poisson distribution having the same mean number of extracted sources.

using Eq. 3, serve as our guiding estimates of the source reliabilities in the absence of prior knowledge of the actual source counts. Figure 14 shows the same information in a slightly different way. Histograms of the extracted source brightnesses are compared to the histograms of the input simulation source brightnesses (when they exist). The ratio of the two histograms yields the reliability that was plotted in Figure 13. The simulations lead us to expect an average of 1.34 false detections in our field if the source counts are similar to the Béthermin et al. (2011) model, or 0.55 sources per field if the sources are more numerous as in the Lapi et al. (2011) model. Finally Figures 15 and 16 show histograms of the total number of sources extracted and histograms of the number of false sources extracted across the 3000 simulations. The histograms tend to show a slight positive tail with respect to Poisson distributions with

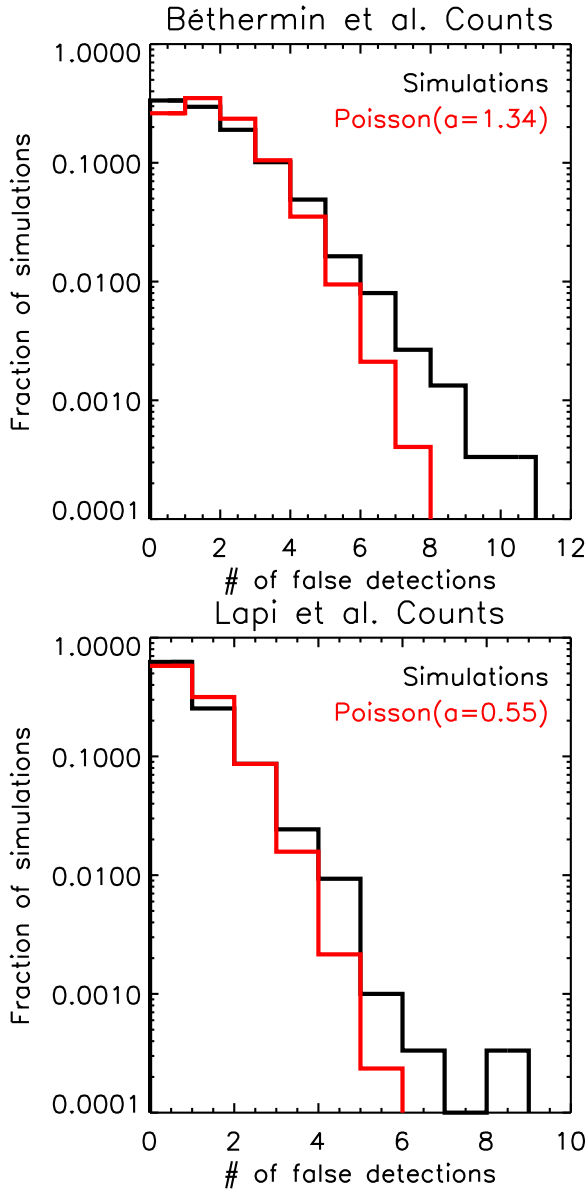


FIG. 16.— Histograms of the number of *false* sources extracted per field, over the set of 3000 simulated fields. Overplotted is the Poisson distribution having the same mean number of false sources.

the same mean number of sources extracted.

5. DISCUSSION

5.1. 2 mm number counts

Figure 17 depicts the observed cumulative number counts, $N(> S)$, as a function of the deboosted 2 mm flux density, S , compared with the predicted galaxy counts in Béthermin et al. (2011) (the solid line is a power-law interpolation between their 2.1 and 1.3 mm models), and the model from Lapi et al. (2011) (dashed line). These counts are not binned due to the small number of sources. Instead, as in e.g. Borys et al. (2003), we plot the number of sources at each deboosted flux density, divided by the effective area for the detection of sources of a given flux density. Following Coppin et al. (2006), the effective area is calculated as the product of the maximum area of field and the fit to the completeness function (Fig.

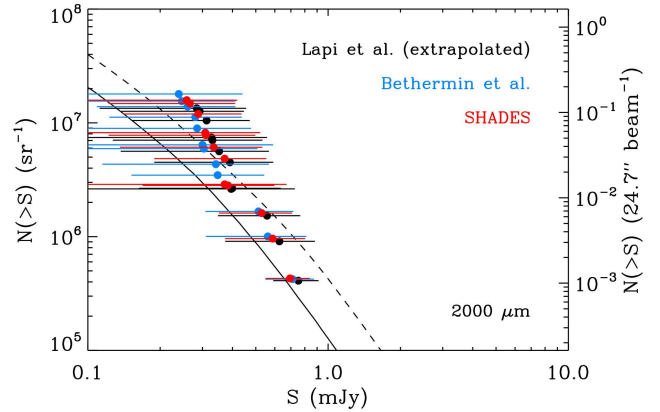


FIG. 17.— The number counts $N(> S)$ as a function of the deboosted 2mm fluxes (S). The black symbols show the deboosting using the Lapi et al. (2011) model extrapolated down to 0.01 mJy. The blue and red show the deboosted number counts when the Béthermin et al. (2011) and the SHADES (Coppin et al. 2006) models are used to calculate the deboosting. The solid line is an interpolation of the 1.38 and 2.1 mm model number counts of Béthermin et al. (2011), and the dashed line is the model from Lapi et al. (2011). The right-hand axis labels the counts in terms of number of sources per beam for convenience in assessing confusion. The counts were corrected for the signal-to-noise dependent effective area for source extraction in the map, as well as for the expected number of false detections from Table 1.

10), although in our case the functional form is simpler and involves only one free parameter. The effective area ranges from 0.21 to 0.90 times the maximum area.

Instead of plotting the cumulative number counts as a simple stair-stepped line, we include the uncertainties for each of the deboosted flux densities. Figure 17 also demonstrates that the cumulative number counts are independent of the three deboosting models were used, considering the uncertainties in the deboosted flux densities. We note that the counts are somewhat steeper than the models.

On a cautionary note, our number counts represent measurements from a small patch of the sky, therefore allowing for quite some uncertainty in terms of cosmic variance, which in the context of AzTEC 1.1 mm number counts is being discussed in Scott et al. (2012).

5.2. Source properties and associations

The GISMO 2-mm datum typically adds an important long-wavelength constraint to the FIR Spectral-Energy Distribution (SED) of a distant ($z > 2$) infrared galaxy, provided it is measured with sufficient precision. When it is the only datum on the Rayleigh-Jeans side of the thermal greybody spectrum, it provides a critical constraint on the infrared luminosity of the galaxy and the cold star-forming dust mass. And, in combination with other Rayleigh-Jeans data (such as 850 μm or 1.2 mm), the 2-mm point can furthermore constrain the temperature and the dust emissivity index β , which is a diagnostic of the physical geometry of the dust grains (see e.g. Yang et al. 2007).

Because the large statistical uncertainties associated with the underlying source fluxes in case of the low signal-to-noise ($\text{SNR} < 5$) detections, where statistical flux boosting is significant (section 4.2), we cannot provide accurate 2-mm photometry individually for most of the detected sources, except for GDF-2000.1 and GDF-

2000.3. However, we can use the GISMO data collectively to determine useful constraints for the population as a whole.

To fit the radio-to-FIR SEDs of our sources, we relied on the analytic temperature-distribution models of Kovács et al. (2010), which assumes a power law distribution ($dM_d/dT \propto T^{-\gamma}$) of dust components above a dominant cold-temperature component at T_c . For the fitting we have assumed a characteristic emission diameter of 2 kpc, typical to SMGs, and a dust emissivity index of $\beta=1.5$, typical for starbursts (Kovács et al. 2006, 2010). We use all available continuum data at wavelengths longer than $15 \mu\text{m}$ (rest frame $3 \mu\text{m}$), a regime dominated by thermal dust (millimeter wavelengths to FIR) and synchrotron radiation (at the radio wavelengths). We assume a 10% calibration uncertainty for all measured bands, added in quadrature to the reported detection uncertainties.

The collective fit to all GISMO sources with sufficient photometry (sources 1, 3, 6, 8, 9, 11, 13 and 14) yields $\gamma=7.24 \pm 0.23$, in excellent agreement with local starburst galaxies (see Kovács et al. 2010). Therefore, for the two most significant individual sources, which we discuss in the following, we fix $\gamma=7.2$, and fit only the cold-component temperature, dust mass and the radio-FIR correlation constant q_L (Kovács et al. 2010) or q_{IR} (Ivison et al. 2010). For the synchrotron spectral index ($S_\nu \propto \nu^{-\alpha}$) we assume $\alpha=0.75$.

5.2.1. GDF-2000.1 & GDF-2000.3

GDF-2000.1 and GDF-2000.3 are the two most significant GDF detections with counterparts. Our SED fits for these sources follow the method described in Kovács et al. (2010). The plots show IRAC data (at $\lambda < 8 \mu\text{m}$), but those were not used for the fit. A summary of the fitted parameters is given in Table 6.

The detection of GDF-2000.1, has a signal to noise ratio $\sigma > 5$ and a statistical detection confidence level of 100%. It is associated with AzGN03 (also dubbed: GN1200.02, GN850.10, MM J123633+621407, SMM J123633.8+621408), and GN10 in Pope et al. (2005) and Pope et al. (2006). This is the source discussed in, e.g., Dannerbauer et al. (2008), and in Daddi et al. (2009) who determine its redshift as $z = 4.04$ through measurements of CO(4-3), the redshift we assume for this source. We note that Wirth et al. (2004) report a possible counterpart at $z = 1.34476$, flagged as “very secure z ”, with $> 99\%$ confidence, centered $\sim 3''$ north of the nominal center of GDF2000.1 coordinates. Fig. 18 shows the SED of GDF-2000.1, using all available flux information of this source at other wavelengths.

GDF-2000.3, is detected with a signal to noise ratio of $\sigma > 4$. With 99% the confidence level for detection is very high. GDF-2000.3 has the SCUBA counterpart GN850.39 and MAMBO/AzTEC counterpart AzGN07. The (sub)millimeter counterparts are described in Pope et al. (2006) and Greve et al. (2008). Two sources with known redshifts are positionally consistent with our measured position: GOODS J123711.98+621325.7 with $z = 1.992$, and SMM J123711.9+621331, with $z = 1.990$. Since there are no other IR sources similarly close we consider those plausible counterparts and use the redshift for our SED fit (Fig. 19).

GDF-2000.1 and GDF-2000.3 are HLIRGs, with $L >$

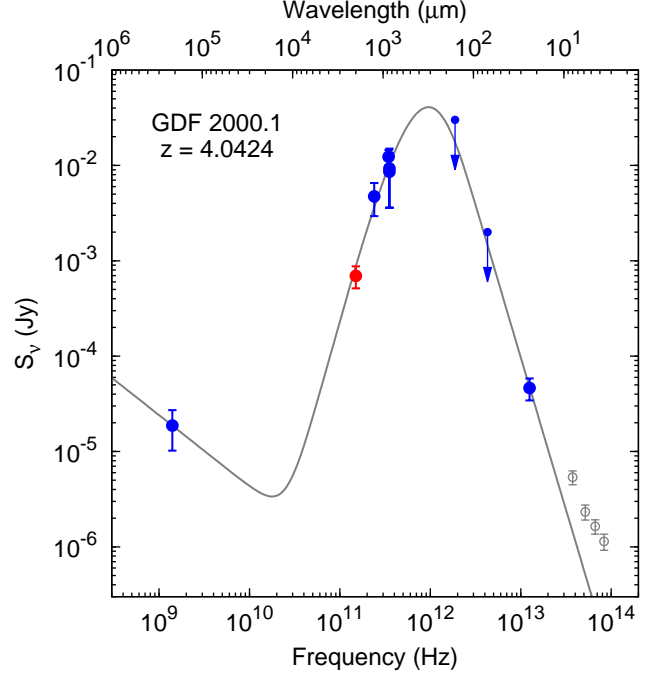


FIG. 18.— SED of GDF-2000.1. The red point at 2mm is the GISMO measurement. The solid line is the model fit to the SED (see text and Table 6), but does not include the 3.6 - $8 \mu\text{m}$ IRAC data as constraints.

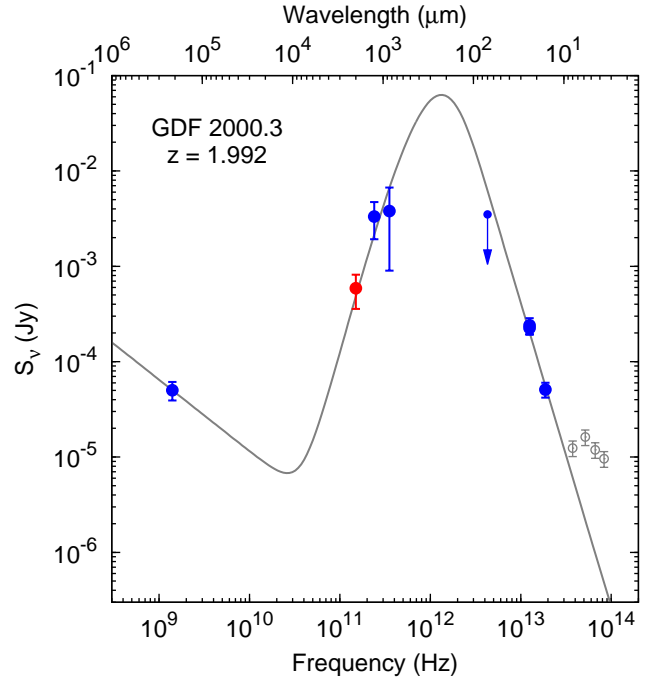


FIG. 19.— SED of GDF-2000.3. Symbols and lines as in Fig. 18.

$10^{13} L_{\text{sun}}$, and $SFR > 1000 M_{\text{sun}}/\text{yr}$. The estimated optical depths (~ 1) at the peak of the emission are typical to SMGs. Both q values are significantly higher than the radio-FIR correlation for SMGs (Kovács et al. (2006a) and Kovács et al. (2010) both found $\langle q_L \rangle \sim 2.13$ for SMGs, with a scatter of 0.12 dex only). Thus, the excess far-infrared emission might need to be explained by the presence of an additional significant heating source besides stars, possibly a powerful AGN. This is unlike

the bulk of the known SMG population, where AGNs, although often present, are not significant contributors (<20%) to the infrared emission.

5.2.2. GDF-2000.6

GDF2000.6 is HDF850.1 ($z = 5.2$), the most prominent among the first ever identified SCUBA Deep Field sources (Hughes et al. 1998). The observed GISMO position and the deboosted 2 mm flux are consistent with the data published in Walter et al. (2012).

5.2.3. Other GDF sources with counterparts

We identify AzGN07 as the counterpart for GDF-2000.8, AzGN08 as the counterpart for GDF-2000.11, and GN850.21 and GN1200.29 as counterparts for the low S/N source GDF-2000.13. We associate CXO J123627.53+621218.0 (with a photometric redshift $z = 4.69$) and AzGN10 as counterparts of GDF-2000.14, since the positions of these two sources are essentially identical, and in very good agreement with the GISMO detection. Furthermore, the high photometric redshift value for the CXO source plus the observed 1 mm flux makes a 2 mm detection at our sensitivity level very likely. Finally, we identify GN850.28 as counterpart to GDF-2000.15.

6. SUMMARY AND CONCLUSION

We have obtained a $7'$ diameter 2 mm continuum deep field map centered on the HDF. The rms in the inner part of the map is $\sim 135\mu\text{Jy/beam}$. The noise in the un-smoothed data very closely follows a Gaussian distribution, indicating its random nature and validating probabilistic source extraction statistics.

We detect 12 sources plus 3 false, negative “detections”. A statistical analysis of the data predicts 2 false detections.

5 of the detected 12 sources have known (sub)millimeter counterparts, including HDF850.1, the first submillimeter galaxy detected by SCUBA. Three more low signal to noise sources have been identified through cross correlation with existing (sub)mm data. The mean redshift of all 7 of the counterparts with known redshifts is $\bar{z} = 3.3$, the median redshift of those sources, which at this low number of sources is probably a better estimate, is $\tilde{z} = 2.91 \pm 0.94$.

Of the remaining 7 detected sources which have no (sub)millimeter counterpart, statistically we expect 5 to be real.

The jackknife test of the smoothed data shows an almost perfect Gaussian distribution for the S/N histogram. The S/N histogram of the normally processed, smoothed data shows a clear excess beyond a Gaussian distribution, which mostly can be contributed to 12 astronomical sources in the field. A small symmetric excess remains after the resolved sources are subtracted from the image. This most likely indicates the presence of confusion noise in our data.

We would like to thank Carsten Kramer, Santiago Navarro, David John, Albrecht Sievers, and the entire IRAM Granada staff for their support during the instrument installation and observations. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain). This work was supported through NSF ATI grants 1020981 and 1106284.

Facility: IRAM:30m (GISMO)

REFERENCES

- Allen, C. A., Abrahams, J., Benford, D. J., et al. 2006, *Proc. SPIE*, 6275
- Allen, C. A., Benford, D. J., Miller, T. M., et al. 2008, *Journal of Low Temperature Physics*, 151, 266
- Baars, J. W. M., Hooghoudt, B. G., Mezger, P. G., & de Jonge, M. J. 1987, *A&A*, 175, 319
- Benford, D. J., Staguhn, J. G., Allen, C. A., & Sharp, E. H. 2008, *Proc. SPIE*, 7020, 56
- B  thermin, M., Dole, H., Lagache, G., Le Borgne, D., & Penin, A. 2011, *A&A*, 529, A4
- Blain, A. W., Ivison, R. J., & Smail, I. 1998, *MNRAS*, 296, L29
- Borys, C., Chapman, S., Halpern, M., & Scott, D. 2003, *MNRAS*, 344, 385
- Borys, C., Scott, D., Chapman, S., et al. 2004, *MNRAS*, 355, 485
- Brandt, N. 2008, *AAS/High Energy Astrophysics Division*, v. 10, pp. 32.01- “Recent Results from the Deepest Chandra Surveys: Adventuring Through the Distant X-ray Universe”
- Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, *Nature*, 470, 233
- Carlstrom, J. E., Ade, P. A. R., Aird, K. A., et al. 2011, *PASP*, 123, 568
- Condon, J. J. 1974, *ApJ*, 188, 279
- Coppin, K., Halpern, M., Scott, D., Borys, C., Chapman, S. 2005, *MNRAS*, 357, 1022
- Coppin, K., Chapin, E. L., Mortier, A. M. J., et al. 2006, *MNRAS*, 372, 1621
- Coppin, K., Halpern, M., Scott, D., et al. 2008, *MNRAS*, 384, 1597
- Cowie, L. L., Barger, A. J., Hu, E. M., Capak, P., & Songaila, A. 2004, *AJ*, 127, 3137
- Crawford, T. M., Switzer, E. R., Holzapfel, W. L., et al. 2010, *ApJ*, 718, 513
- Daddi, E., Dannerbauer, H., Krips, M., et al. 2009, *ApJ*, 695, L176
- Dannerbauer, H., Walter, F., & Morrison, G. 2008, *ApJ*, 673, L127
- Dobbs, M., Halverson, N. W., Ade, P. A. R., et al. 2006, *New Astronomy Reviews*, 50, 960
- Frayer, D. T., Huynh, M. T., Chary, R., et al. 2006, *ApJ*, 647, L9
- Gialalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, 600, L93
- Greve, A., Panis, J.-F., Thum, C. 1996, *A&A*, 115, 379
- Greve, T. R., Pope, A., Scott, D., et al. 2008, *MNRAS*, 389, 1489
- Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, *Nature*, 394, 241
- Irwin, K. D., Vale, L. R., Bergren, N. E., et al. 2002, *Low Temperature Detectors*, 605, 301
- Ivison, R. J., Alexander, D. M., Biggs, A. D., et al. 2010, *MNRAS*, 402, 245
- Kov  cs, A., Chapman, S. C., Dowell, C. D., et al. 2006a, *ApJ*, 650, 592
- Kov  cs, A. 2006b, PhD thesis, Caltech
- Kov  cs, A. 2008, *Proc. SPIE*, 7020, 45
- Kov  cs, A., Omont, A., Beelen, A., et al. 2010, *ApJ*, 717, 29
- Lapi, A., Gonz  lez-Nuevo, J., Fan, L., et al. 2011, *ApJ*, 742, 24
- Magnelli, B., Lutz, D., Santini, P., et al. 2012, *A&A*, 539, A155
- Morrison, G. E., Owen, F. N., Dickinson, M., Ivison, R. J., & Ibar, E. 2010, *ApJS*, 188, 178
- Negrello, M., Magliocchetti, M., Moscardini, L., et al. 2004, *MNRAS*, 352, 493
- Penner, K., Pope, A., Chapin, E. L., et al. 2011, *MNRAS*, 410, 2749
- Perera, T. A., Chapin, E. L., Austermann, J. E., et al. 2008, *MNRAS*, 391, 1227
- Pope, A., Borys, C., Scott, D., et al. 2005, *MNRAS*, 358, 149
- Pope, A., Scott, D., Dickinson, M., et al. 2006, *MNRAS*, 370, 1185
- Rodighiero, G., Lari, C., Pozzi, F., et al. 2006, *MNRAS*, 371, 1891

- Sargent, M. T., Béthermin, M., Daddi, E., & Elbaz, D. 2012, *ApJ*, 747, L31
- Scott, K. S., Wilson, G. W., Aretxaga, I., et al. 2012, *MNRAS*, 423, 575
- Sharp, E. H., Benford, D. J., Fixsen, D. J., et al. 2012, *Proc. SPIE*, 8452, 3
- Sharp, E. H., Benford, D. J., Fixsen, D. J., et al. 2008, *Proc. SPIE*, 7020, 66
- Staguhn, J. G., Benford, D. J., Allen, C. A., et al. 2006, *Proc. SPIE*, 6275, 62751D
- Staguhn, J. G., Benford, D. J., Allen, C. A., et al. 2008, *Proc. SPIE*, 7020, 3S
- Staguhn, J. G., Benford, D. J., Fixsen, D. J., et al. 2012, *Proc. SPIE*, 8452
- Swetz, D. S., Ade, P. A. R., Amiri, M., et al. 2011, *ApJS*, 194, 41
- Teplitz, H. I., Siana, B., Brown, T. M., et al. 2006, *AJ*, 132, 853
- Walter, F., Decarli, R., Carilli, C., et al. 2012, *Nature*, 486, 233
- Wei, A., Kovács, A., Coppin, K., et al. 2009, *ApJ*, 707, 1201
- Wirth, G. D., Willmer, C. N. A., Amico, P., et al. 2004, *AJ*, 127, 3121
- Yan, H., Dickinson, M., Stern, D., et al. 2006, *New Astronomy Reviews*, 50, 127

APPENDIX

Figure 20 shows the posterior probability distributions for all GISMO sources, indicating the Bayesian probability densities that the detected source flux arises from an intrinsic source flux of S_i . The probabilities account for up to 3 overlapping resolved sources contributing to the observed flux, and take confusion at the faint end into account. Since our deboosting method is based on resolved sources only, no additional zero-level adjustment is necessary. As a result, the distribution naturally does not extend to negative fluxes, as is evident in the figure. The distributions are shown for the number count models of Béthermin et al. (2011) in red/solid, the extended Lapi et al. (2011) counts in blue/dashed, and the scaled broken powerlaw SHADES counts in cyan/dashed-dotted.

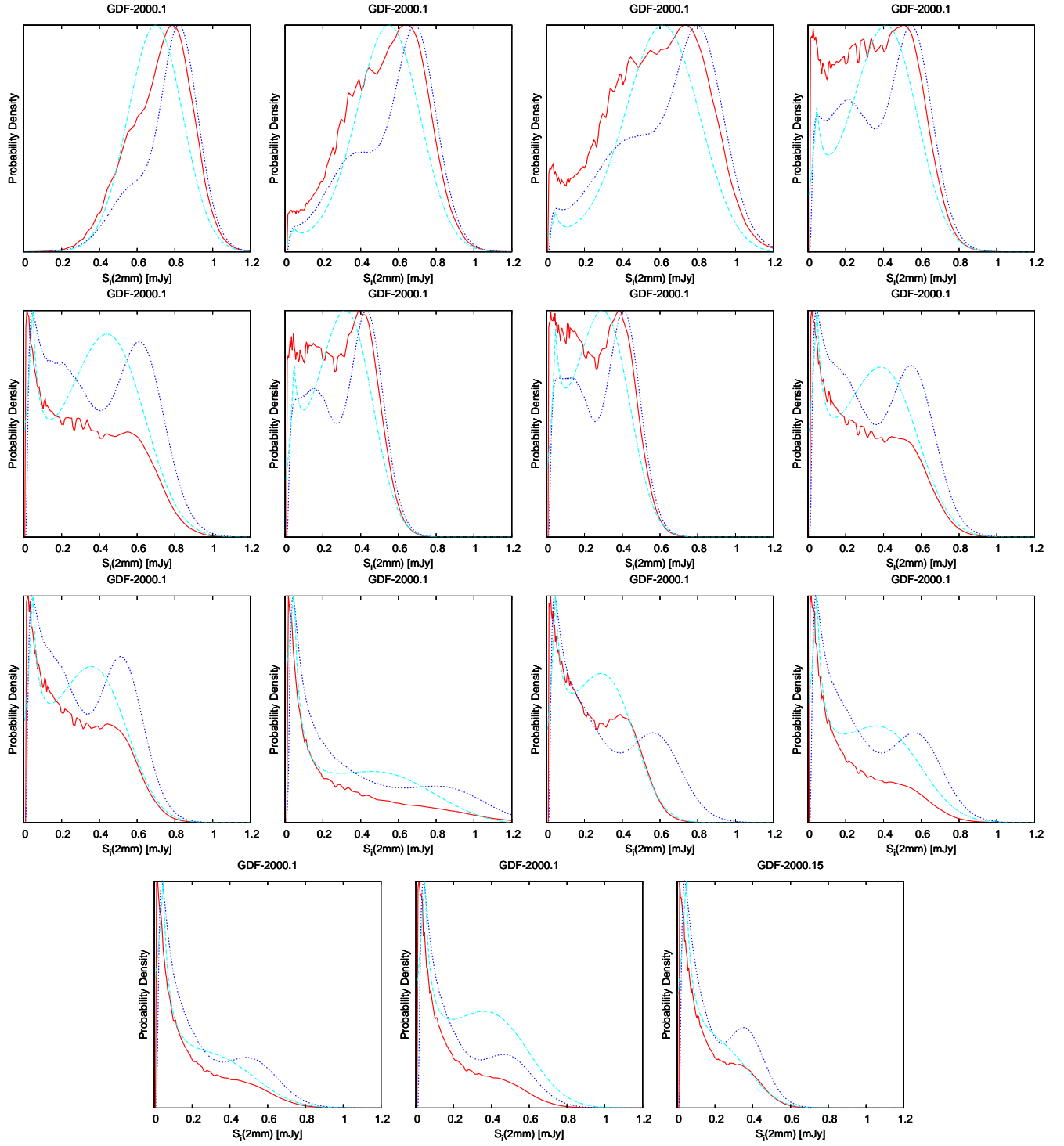


FIG. 20.— Posterior probability distributions for all GISMO sources, indicating the Bayesian probability densities that the detected source flux arises from an intrinsic source flux of S_i . The probabilities account for up to 3 overlapping resolved sources contributing to the observed flux, and take confusion at the faint end into account. The distributions are shown for the number count models of Béthermin et al. (2011) in red/solid, the extended Lapi et al. (2011) counts in blue/dashed, and the scaled broken powerlaw SHADES counts in cyan/dashed-dotted. Note, that the jagged curves resulting from the Béthermin et al. (2011) counts are not a property of our algorithm, but are inherent to the input counts model.